Establishing and Validation of Modified CRE/Loxp System Useful In Temporal and Spatial Control of Genetic Knocking Out Using In Vitro Approaches by Induced Alpha Complementation

Ahmed M Hamdan1,2*, Mohammed M H Al-Gayyar3,4 and Abdullah A Alyoussef5

1Department of Genes and Behavior, Max-Planck Institute for Biophysical Chemistry, Gottingen, Kingdom of Saudi Arabia
2Department of Pharmaceutics, Faculty of Pharmacy, University of Tabuk, Tabuk, Kingdom of Saudi Arabia
3Department of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Tabuk, Tabuk, Kingdom of Saudi Arabia
4Department of Clinical Biochemistry, Faculty of Pharmacy, University of Mansoura, Mansoura, Egypt
5Department of Internal Medicine, Faculty of Medicine, University of Tabuk, Tabuk, Kingdom of Saudi Arabia

Abstract

DNA recombinases are used to manipulate the genomic structure and to control the genetic expression in all living organisms. Cre is a P1 bacteriophage-derived tyrosine recombinase mediating the site-specific recombination between two loxP DNA recognition sites. Cre/loxP system allows generation of tissue specific mutations and is widely used in bioengineering and in mouse genetics. However, various factors limit its applicability such as lack of temporal control over its recombination activity, presence of cellular compensatory mechanisms, loss of site-specificity at high expression levels and its limited use for conditional recombination in certain brain structures due to a lack of sufficiently selective promoters. One approach used to overcome these drawbacks is the so-called split-iCRE technique that employs complementation of split-iCRE fragments via artificial FKBP12-rapamycin or α-helix interactions. Rapamycin is a pharmacologically active substance while leucine zipper-mediated complementation so far only insufficiently restores recombinase activity. Here we introduce a system for controlling α-complementation of two independent iCRE fragments under the control of two different promoters to reconstitute recombinase activity. Unsplit-iCRE was split between Lys130 and Asp132. Fragments were complemented restoring ~95% of the recombinase activity with very little background activity. Adding an external nuclear localization signal to the C-terminal fragment resulted in even higher enzymatic activity. Using an extended rigid linker between the polypeptide and the yeast GCN4-coil/coil leucine zipper domains was more efficient than a semi-flexible separator. The system was validated by knocking out the essential circadian clock component Bmal1 (Arntl) in MEF cells resulting in a loss of clock function in MEF cells.

Keywords: Complementation; Conditional Knocking Out; Leucine Zipper; Recombinase; Split-iCRE

Abbreviations: Bmal1: Brain and Muscle Arnt-Like Protein 1; cCre: C-Terminal Fragment of Split-iCRE; nCRE: N-Terminal Fragment of Split-iCRE; DEX: Dexamethasone; DMEM: Dulbecco Modified Eagle's Medium, NLS: Nuclear Localization Signal, MEF: Mouse Embryo Fibroblast; SCN: Supra Chiasmatic Nucleus.
Introduction

All organisms show both endogenous physiological and behavioral daily variations of approximately 24 hr. This variation enables the organism to adapt daily environmental alteration. In mammals, circadian clock is generated centrally within the Supra Chiasmatic Nucleus (SCN). Such oscillation maintains proper phase alignment of peripheral tissue clocks ubiquitously distributed in nearly all cells. Thus, the brain SCN clock provides “standard time” for all peripheral tissue clocks [1]. Disturbance of the molecular clock is the main cause of many diseases such as mood disorders, rheumatoid arthritis and cancer [2]. The molecular mechanism that controls this daily variation includes transcriptional-translational negative feedback loops. CLOCK–BMAL1 heterodimer binds to an E-box of the promoter region activating the transcription of Period (Per1, Per2, and Per3) and Cryochrome (Cry1 and Cry2). This transcriptional up-regulation is down-regulated through PER and CRY by direct protein–protein binding interacting with CLOCK–BMAL1 heterodimer [3]. Genetic manipulation of clock genes in the SCN to model some diseases has been challenging.

Cre/loxP system is derived from the tyrosine site-specific recombinase system of the bacterial P1 phage [4]. In biomedical research, it is extensively used for spatial and temporal gene recombination. It consists of two major parts. First, the unsplit-iCRE recombinase enzyme catalyzes deletion, inversion, integration or translocation of the genetic material. Second, two consensus sequences placed in the same orientation on the same DNA strand called loxP mediate iCRE targeting to specific DNA loci. Each loxP motif consists of 34 BP [5-7]. Owing to its simplicity, efficiency and economy, Cre/loxP system is widely used in bioengineering, e.g. for studying gene function by functional deletion. Nevertheless, the standard Cre/loxP approach contains several limitations. First, it lacks temporal control over the recombination event. As a consequence, Cre/loxP-mediated deletion of an ontogenetically important gene in mice may lead to death at early embryonic stages, thus hindering the study of the genetic function in other physiological processes at later stages. Second, the presence of cellular compensatory mechanisms may mask recombination activity at low expression levels. Third, at high and/or prolonged expression unsplit-iCRE can lose its site-specificity [8]. This may lead to interaction between unsplit-iCRE recombinase and loxP-like sequences naturally present in the mammalian genome, causing uncontrolled genetic rearrangements and disturbance of cellular physiology [9]. Fourth, already the insertion of loxP sites may affect normal gene expression. In consequence, numerous publications may actually not use the correct control group for their experiments. Finally, due to absence of tissue-specific selective promoters, usage of unsplit-iCRE is limited for conditional recombination in certain brain nuclei [10-13]. In consequence, exploring strategies to optimize the Cre/loxP system thus remains a current point of research. Four principal approaches have been deployed to overcome the above-mentioned problems. (i) unsplit-iCRE recombinase can be expressed using cell- or tissue-specific promoters to achieve better spatial control. However, it is not always possible to find a promoter specifically expressed at sufficient levels in the desired cell type or tissue. This is particularly true for complex organs such as the brain.[14] (ii) Transcription and, thus, activity of unsplit-iCRE recombinase can be regulated by using an inducible promoter sensitive to, e.g., tetracycline or interferon-induced promoters to achieve temporal control [15-17]. However, both tetracycline and interferon are pharmacologically active compounds affecting cellular physiology. Another approach uses a fusion of unsplit-iCRE and a steroid receptor binding/nuclear translocation site. This approach allows post-translational regulation of unsplit-iCRE nuclear targeting by applying synthetic steroids such as Tamoxifen or RU 486 [18-21]. However, these systems are often not tight enough,
resulting in background recombination effects under non-stimulated conditions [22]. (iii) Another approach uses programmed stem cells for temporal control of unsplit-iCRE recombinase activity [23, 24]. (iv) Finally, the structure of the Cre recombinase can be modulated to improve temporal and spatial control of its activity. The split-iCRE approach uses forced non-covalent association of differentially expressed N- and C-terminal parts of unsplit-iCRE fused to immunophilin FK509-binding protein and FKBP12-rapamycin-associated protein, respectively, by rapamycin or natural alpha-helix protein-protein interaction [22-25]. Rapamycin is an immunosuppressive agent. It can affect both cellular viability and overall physiology. Other trials used synthetic leucine zipper-induced complementation of split-iCRE fragments, but this approach only restores little unsplit-iCRE recombinase activity (2-3% compared to non-split approaches) [26, 27].

In the present paper, we explored a new strategy for improving tempo-spatial regulation of Cre using non-covalent alpha-complementation of two inactive split-iCRE fragments mediated by fusion to yeast GCN4-coil/coil domains linked to the split-iCRE fragments by sterically rigid linkers. Each fragment can be expressed under the control of an independent promoter. For validation, we used CMV and PGK promoters to study the different kinetics of each fragment. CMV promoter is about 3 times stronger than PGK promoter [28]. Using the online Protein Structure Visualization tools:

http://www.expasy.org/proteomics/protein_structure,
http://www.pdb.org/pdb/home/home.do,
http://ligin.weizmann.ac.il/space, we studied the Cre 3D and 2D structures. So, we choose splitting site between Lysine 130 and Asparagine 132. We assumed that this splitting site will not disturb the C-shaped clamp of Cre surrounding the DNA. It will also keep the previously published active sites of unsplit-iCRE intact (29). It will also divide unsplit-iCRE between the two major α-helices; nCre of 13 kDa and cCre of 25.5 kDa [5, 29].

Moreover, the reconstituted fragments will be perpendicular to the planner Holliday junction avoiding any distortion of the unsplit-iCRE structure after complementation using yeast GCN4-coil/coil leucine zipper [19, 30]. We removed Glutamine 131 amino acid to avoid the imposition of directionality on the linker leaving a space (~13.89 Å) for the linker. It has been previously reported that there is an intrinsic NLS signal in the proximal terminal part of the C-fragment of unsplit-iCRE [29, 31]. So, we tested the effect of adding NLS to either the N-terminal or the C-terminal part of the C-fragment. We tested also the effect of removing the external NLS from the C-fragment. We found that, our system succeeded to restore ~95% of the recombinase activity after reconstitution of fragments. There was kinetic difference between these two fragments indicating that it was preferably to add C-fragment to the stronger promoter. Moreover, addition of NLS was essential for obtaining high recombinase activity. Removing of the NLS from the C-fragment decreased the resultant recombinase activity by ~65%. Reconstitution of the recombinase activity was monitored through protein-protein interactions. Each split-iCRE fragment was fused to yeast GCN4-coil/coil domain to force dimerization. Linker design is an important step to efficiently separate between active domains in the bifunctional protein. It has been published to use semi-flexible linker [22-30]. Here, we found that rigid linker was better than semi-flexible one (about 1.3 fold). The mechanism of the recombinase activity of unsplit-iCRE involves the construction of a highly ordered protein-DNA complex between the two loxP sequences and four Cre molecules interacting with each other. This step is followed by the formation of Holliday junction and a succession of cleavage-relegation steps [32]. Long rigid linker may keep the configuration of the unsplit-iCRE after reconstitution reducing the steric hindrance for DNA binding. Testing the reconstituted recombinase activity on mammalian genome, we tested the effectiveness of this approach in disrupting circadian clock function in MEF cells derived from Bmal1<sup>−/−</sup> mice. Dexamethazone treated MEF cells from Bmal1<sup>−/−</sup> mice gave the same rhythm disruption profile of Clock, Per1, Per2, Cry2, Dec1 and Rev-erb-a of Bmal1<sup>−/−</sup> mice. Also CLOCK protein was present in Bmal1<sup>−/−</sup> MEF cells in the cytoplasm not in the nucleus after transfection with both nCre and cCre plasmid vectors and dexamethazone stimulation previously published for Bmal1<sup>−/−</sup> MEFs [33]. Despite the
complex theory of ligand-induced complementation, we show that it can efficiently in vitro regulate the recombinase activity of unsplit-iCRE and it can be promisingly an alternative approach for the regulation of its both temporal and special activity as well. This system opens the possibility to study gene function efficiently by enabling conditional knocking out especially for Neuroscience.

Materials and Methods

Construction of expression plasmid vectors

The sequences coding for the fusion proteins were constructed by PCR amplification. We used the humanized version of unsplit-iCRE recombinase using the pNLS-iCRE cDNA (with nuclear localization signal present in the Cre’s C-terminus; a kind gift from Dr. Xunlei Zhou, University of Heidelberg, Germany). One terminal cDNA ORF (called nCre) encodes the N-terminal part of unsplit-iCRE (aa 1-130) fused to a nuclear localization signal (NLS; PKKKRKVPKKKRKV) was inserted into the SacI and BamHI restriction sites of pCMV-Tag-4A (Stratagene, La Jolla, CA, USA). The other cDNA ORF (called cCre) encodes a protein of the C-terminal fraction of unsplit-iCRE (aa 132-345) fused to a NLS was inserted into the SacI and BamHI restriction sites of pCMV-Tag-3A (Stratagene, La Jolla, CA, USA). cDNA encoding the leucine zipper domain of yeast GCN4 coil/coil (a kind gift from Dr. Johannes Hirrlinger, Faculty of Medicine, and Interdisciplinary Centre for Clinical Research, University of Leipzig, Germany) was PCR-amplified with a semi-rigid linker (ASPSNPGASNGS) or a rigid linker (ASGAEAAKAGGGS) and inserted into the EcoRI and EcoRV restriction sites of each vector. Linker sequences have been published [22, 30]. The GCN4 coil/coil domain will be called NZ and CZ for the plasmids expressing nCre and cCre respectively. The final expression constructs were iCre (1-130)-NLS-Linker-NZ-FLAG or iCre (132-345)-NLS-Linker-CZ-MYC. The resulting fusion proteins were named nCre and cCre, respectively (Figure: I A-C). Primer sequences are shown in Table 1.

<table>
<thead>
<tr>
<th>Amplicon</th>
<th>Sequence of the primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsplits-iCRE</td>
<td>5'-CCCGAGCTCGCCATGATGTCCTGGACCAAAGAAGAGG -3'</td>
</tr>
<tr>
<td></td>
<td>/5'-CGGGATCCATGTCCTGGACCAAAGAAGAGG</td>
</tr>
<tr>
<td></td>
<td>ACCTTTCTTTTTTTTGGATGTGACGTTCATCCCCATCTCGAGCA-3'</td>
</tr>
<tr>
<td>nCre</td>
<td>5'-CCCGAGCTCGCCATGATGTCCTGGACCAAAGAAGAGG</td>
</tr>
<tr>
<td></td>
<td>/5'-CGGGATCCATGTCCTGGACCAAAGAAGAGG</td>
</tr>
<tr>
<td></td>
<td>TTTCACACACATTTTGTGAGCTGTCATGGCCAGGTGTGAGCAGCACAGGAG-3'</td>
</tr>
<tr>
<td>cCre</td>
<td>5'-CCCGAGCTCGCCATGATGTCCTGGACCAAAGAAGAGG</td>
</tr>
<tr>
<td></td>
<td>/5'-CGGGATCCATGTCCTGGACCAAAGAAGAGG</td>
</tr>
<tr>
<td></td>
<td>TTTCACACACATTTTGTGAGCTGTCATGGCCAGGTGTGAGCAGCACAGGAG-3'</td>
</tr>
</tbody>
</table>

Table 1: List of primer sequences used for construction of expression plasmid vectors.
Cell culture techniques

All cell culture materials were obtained from Invitrogen. HEK 293-T/17 cells were obtained from ATCC (CRL-11268, American Tissue Culture Collection; Rockville, MD). HEK293-T/17 cells were grown under standard tissue culture conditions. Cells were co-transfected with a mixture of pCMV-STOP-loxP-STOP-Luc and iCre(1-345)-NLS-FLAG, nCre and/or cCre, or empty pCMV-Tag-4A vector using FuGENEHD transfection reagent (Promega, Mannheim, Germany) using the manufacturer’s protocol. Immortalized Mouse Embryonic Fibroblasts (MEFs) prepared from Bmal1fl/fl, Bmal1^-/- [33] or congenic C57BL/6 Wild-Type (WT) mice in Dulbecco’s Modified Eagle’s Medium (DMEM; Sigma-Aldrich St. Louis, MO, USA) supplemented with 10% FBS (AFC Biosciences, Lenexa, USA) at 37 °C under a humidified 5% CO₂ atmosphere for . The culture medium for Bmal1fl/fl, Bmal1^-/- or WT MEFs was supplemented with 20 µM β-mercaptoethanol and 1x non-essential amino acid mix. For circadian synchronization, cells were incubated with 10 µM forskolin (FSK; Sigma-Aldrich), 50% Fetal Bovine Serum (FBS) or 10 nM Dexamethasone (DEX) (Sigma-Aldrich) for 2 h. To assess the circadian oscillation of clock genes expression, semi-confluent cultured MEFs were stimulated with 10 nM DEX for 2 h; then the medium was replaced with DMEM supplemented with 10% FBS. Cells were harvested for RNA extraction at the indicated times after DEX stimulation. Total RNA was extracted using RNAiso (Takara Bio, Otsu, Japan), and mRNA levels were measured by quantitative RT-PCR.

Stable cell lines

Immortalized Bmal1^-/^- and Bmal1^-/- MEFs were transfected with plasmid vectors expressing nCre and cCre by FuGENEHD Transfection Reagent (Promega, Mannheim, Germany) using the manufacturer’s protocol. Clones stably expressing the nCre or the cCre constructs for 72 h were isolated using CherryPickercell capture system (Clontech, Saint-Germain-en-Laye, France) according to the manufacturer’s protocol. Stably transfected MEFs were used for experiments after 24-h incubation with DMEM, 20% FBS.

Quantitative RT-PCR

cDNA was synthesized by reverse-transcribing 0.4 µg of total RNA using oligo(dT)15 primers and reverse transcriptase (Invitrogen). The cDNA equivalent of 12 ng of RNA was amplified by PCR in a CFX-96 real-time PCR system using SYBR-GREEN detection (Bio-Rad, München, Germany). Relative quantification was done using the ΔΔCT method with β-actin as reference gene. Sequences for PCR primers are described in Table 2.
Table 2: List of primers used for qRT-PCR

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence of the primer</th>
</tr>
</thead>
</table>
| clock | 5'-AGGCACAGACATTATCCG-3  
       | 5'-ACCGTCTCATCAAGGGAC-3 |
| per1  | 5'-TCTGGTTCGGGATCCACGAA-3  
       | 5'-GAAGAGTCGATGCTGCAAG-3 |
| per2  | 5'-CACCCCTGAAAAGAAGTGCGA-3  
       | 5'-CAACGCAAGGAGCTCAAGT-3 |
| cry2  | 5'-TTGAAGAAGACCCAGAAC-3  
       | 5'-ACTTTCATTCCTCTTC-3 |
| dec1  | 5'-AGATCAACTGCCTGGACAGCATCCTCAG-3  
       | 5'-AAACCTCTGGGTTCCTACCA-3 |
| rev-erbα | 5'-AGGCTTCCGTCGACCTTCTCA-3  
       | 5'-GTCGTGTCCTTACAGTTGAACA-3 |
| β-Actin | 5'-CACACCTTCTACATGAGCTGC-3  
       | 5'-CATGATCTGGGTCATCTTTC-3 |

Luciferase assays

Luciferase assays were performed using a dual-luciferasereporter assay kit (Promega) according to the manufacturer’s protocol. Briefly, we seeded HEK293-T/17 cells at a density of 1x10^5/well in 24-well culture plates. Cells were transfected 18 h later with 100 ng/well of pCMV-loxP-STOP-loxP-Luc reporter vector and 1-2 µg/well of expression vectors using empty pCMV-Tag-4A or pCMV-Tag-5A as controls. The pRL-TK vector (0.5 ng/well; Promega) was also co-transfected as internal control of transfection efficiency. After 24 h, cells were harvested and the cell lysates were analyzed using a dual luciferase reporter assay system (Promega). The ratio of firefly (expressed from the reporter construct) to Renilla (expressed from pRL-TK) luciferase activities in each sample served as a measure of normalized luciferase activity.

Immunoblotting

Harvested cells were incubated on ice-cold RIPA buffer. For nuclear protein extraction, cells were suspended in freshly prepared buffer A (10 mM HEPES and 10 mM KCl in Dist. H2O) and incubated at 4°C for 30 min. After adding NP-40, cells were incubated at 4°C for another 15 min and resuspended by vortexing for 30 sec. Nuclei were pelleted in a microfuge at maximum speed for 20 sec. The supernatant was used as the cytoplasm fraction. Nuclear and cytoplasm fractions were normalized by measuring total protein concentrations using the Bio-Rad De Protein Assay Kit according to the manufacturer's
protocol. Proteins were separated by 6% SDS-PAGE and transferred to Immobilon-P membranes (Millipore). Endogenous CLOCK protein was detected with guinea pig anti-CLOCK antibodies. Anti-H4 histone (N-18) and anti-actin (C-11) antibodies from Santa Cruz Biotechnologies were used to control for nuclear and cytoplasmic protein separation. HRP-conjugated rabbit IgG (Santa Cruz Biotechnologies) or guinea pigs IgG (Jackson Laboratory) were used as secondary antibodies. Proteins were visualized with by ECL detection (Amersham).

Statistical Analysis
For descriptive statistics of quantitative variables, the mean ± standard error was used. Normality of the sample distribution of each continuous variable was tested with the Kolmogorov–Smirnov (K–S) test. One-way analysis of variance (ANOVA) was used to compare means between groups. Once the differences exist among the means, post-hoc Bonferroni correction test was calculated. Statistical computations were done on a personal computer using Microsoft Excel 2007. Statistical significance was predefined as P ≤ 0.05.

Results

Principle approach for splitting unsplit-iCRE
The principle approach for optimizing the splitting of functional proteins is to study the 3D configuration of active domains, functionally and sterically important amino acid distribution, and the transient interaction with substrates, known activators and inhibitors. Such structural analysis of the unsplit-iCRE enzyme used in this study (accession number AY056050) defined a suitable position of splitting between Lys130 and Asp132. This position lies between the two functionally essential inter-helical loops of the unsplit-iCRE. It is fully exposed upward to the external environment. Moreover, it appeared not involved in intra- or intermolecular interactions. EXPASY modeling predicted that splitting at this site would largely preserve the 3D-configuration of subdomains in each isolated fragment (Figure 1B).

Figure 1: Schematic presentation of the Split-Cre Recombinase system.
The enzyme is split into two fragments; nCre and cCre (Orange). (A) Our designed system: nCre codes the N-terminal fragment between amino acids 1-130. cCre codes the C-terminal part between amino acids 132-345 is fused to nuclear localization signal; NLS (Yellow) [53]. This allele is inserted with the Kazak's consensus sequence [54] in the SacI and BamHI restriction sites of pCMV-Tag-4A plasmid vector (with FLAG; Red). cCre codes the C-terminal fragment between amino acids 132-343. The C-terminal part of the unsplit-iCRE fragment is fused to double strength nuclear localization signal; NLS (Yellow). This allele is inserted in the SacI and BamHI restriction sites of pCMV-Tag-5A plasmid vector (with MYC; Green). Sequences coding the Rigid linker (Violet) are fused with the yeast GCN4 coil/coil domain (Blue) and inserted between the EcoRI and EcoRV restriction sites of both pCMV-Tag-4A and pCMV-Tag-5A plasmid vectors (B) Currently used system [30]: nCre codes the N-terminal fragment between amino acids 19-59. cCre codes the C-terminal fragment between amino acids 60-343. (C) Wild-type Cre plasmid vector used as a control fused to pCMV-Tag-4A plasmid vector with 2XNLS, in the SacI and BmaHI restriction sites. Then, we inserted Rigid or semi-flexible linkers fused to the yeast GCN4 coil/coil domain.
in the EcoRI and EcoRV restriction sites. This diagram is drawn to scale as indicated.

Testing the effect of fragments on the restored recombinase activity

A similar split-iCRE approach had been used before, parting the functional protein between amino acids 59 and 60 and using a FKBP12-rapamycin [22] or leucine zipper-induced complementation [25, 30]. Jullien et al. [22, 34] built their strategy on leaving the C-shaped clamp moiety of unsplit-iCRE-recombinase intact to surround the DNA double strand. We split Cre into two moieties; nCre; 1-130 and cCre; 132-345. We compared the reconstituted recombinase activity of the currently used system [30] and of our system using a pCMV-loxP-STOP-loxP-LUC reporter construct in HEK293-T/17 cells. Luciferase activity showed that about 95% of the unsplit-iCRE-recombinase activity was regained using our 1-130/132-345 system compared to intact unsplit-iCRE. In comparison, a 1-59/60-345 split-iCRE [22] resulted in only 60% efficiency compared to unsplit-iCRE (Figure 2). Different splits may result in different transcriptional/translational kinetics, post-translational modifications, association properties, and recombinase activity. To test the relative contributions of the two fragments we analyzed recombination efficiency after transfecting different ratios of both fragments. We found that expression of each of the two fragments alone did not result in significant recombinase activity. Moreover, when expressing both fragments from a strong promoter (CMV), no differential contribution to recombinase activity was observed (Figure 3A).

On the other hand, under the influence of a weaker promoter (PGK, which has ~3 times less activity than CMV in HEK293-T/17 cells [28]), we found that the high cCre expressing vector ratio the high recombinase activity in a dose-dependent manner. Such behavior did not appear if we increase the dose of nCre expressing vector (Figure 3). We also tested if the introduction of the yeast GCN4-coil/coil domain with the split-iCRE fragments would interfere with cellular physiology. We did not observe any alterations in the growth rate or morphology of either HEK293-T/17 or MEF cells after transfection. Moreover, immunocytochemistry showed no interaction between endogenous activating transcription factor family containing leucine zipper domain and transfected yeast GCN4 (Data are not shown). This result confirmed previously published data [34, 35].

Figure 2: Effect of fragmentation on the restored Cre Recombinase activity

Relative Cre Recombinase activity using different fragments of Cre using synthetic construct pCMV-loxP-STOP-loxP-LUC in HEK 293T/17 cells. Unsplit-iCRE is considered as 100%. New system is our developed system using nCre; coding amino acids 1-130 and cCre; coding amino acids 132-343. Currently used system is that used by Jullien et al. [34]. They used nCre; coding amino acids 19-59 and cCre; coding amino acids 60-343. Each value represents the mean +/- S.E. N=3-5 and **p< 0.01.
Dotted line represents normalization of the relative luciferase activity by Renilla expression vector.

Ahmed et al, Figure 3

Figure 3: Relative Cre Recombinase activity restored using different ratios of fragments for complementation

We test the restored recombinase activity using synthetic construct pCMV-loxP-STOP-loxP-LUC in HEK 293T/17 cells. (A) Different ratios of nCre and cCre expressing vectors under the control of CMV promoter. No statistical significance among different combination of 1ng nCre plasmid coding vector and various amounts of cCre coding plasmid vector. Both nCre and cCre have no recombinase activity without the other fragment. (B) Different ratios of nCre and cCre expressing vectors under the control of PGK promoter. There is an increase of the reconstituted Cre recombinase activity with the increase of the amount of cCre expressing vector in a dose-dependent manner. Meanwhile, there is no increase in the reconstituted Cre recombinase activity with the increase of the amount of nCre expressing vector. Each value represents the mean +/- S.E. N=3-5. Dotted line represents normalization of the relative luciferase activity by Renilla expression vector.

Testing the effect of NLS on the restored recombinase activity

It has previously been reported that synthetic unsplit-iCRE contains a Nuclear Localization Signal (NLS) domain in the C-terminal part of the enzyme [36-39]. This domain can interact with the importin machinery to promote the translocation of unsplit-iCRE into the nucleus. To test whether additional insertion of NLS domains would improve nuclear translocation and, thus, recombination efficiency in our split-iCRE approach, we added NLS N- or C-terminally to the cCre construct (Figure 4A). We found that absence of externally added 2XNLS reduces the recombinase activity by about 65% of the wild-type recombinase activity. Addition of NLS to the C-terminal part of the cCre fragment was 40% higher than that added to the N-terminal part of the cCre fragment (Figure 4B). So, it is more preferable to add 2XNLS to the C-terminal part of the cCre fragment.

Ahmed et al, Figure 4

Figure 4: Relative Cre recombinase activity restored using different positions of nuclear localization signals in cCre fragment.

We test the restored recombinase activity using synthetic construct pCMV-loxP-STOP-loxP-LUC in HEK 293T/17 cells. (A) Schematic diagram for the used plasmid vectors. We changed the position of the NLS allele from the N-terminal part of the cCre fragment (indicated by NLS-C) to the C-terminal part of the cCre (indicated by C-NLS). We used cCre plasmid vector coding no 2XNLS allele (indicated without NLS). Each
value represents the mean +/- S.E. N=3-5 and **p< 0.01. Dotted line represents normalization of the relative luciferase activity by Renilla expression vector.

Testing the effect of Linker on the restored recombinase activity

We tested semi-flexible (ASPSNPGASNGS) and rigid (ASGAEAAAKEAAAKAGGGS) linkers. Rigid linker was better than semi-flexible linker (~1.33 fold) (Figure 5). This may be due to overcoming the steric hindrance occurred during Cre-loxP binding. It also may keep the configuration of the C-Shaped clamp around the DNA in a plane perpendicular to the plan of the DNA. The longer linker used the higher reconstituted recombinase activity.

Disturbed rhythmic expression of Clock genes in Bmal1<sup>−/−</sup> MEFs after transfection with both nCre and cCre

Bmal1 is one the main core-clock gene components [37]. After DEX stimulation, Bmal1<sup>−/−</sup> MEFs showed decrease in the intensity of the expression level of per1, per2, cry2, dec1 and rev-reb with damped rhythmic expression or totally absent of the 24 hr variation than Wild-type MEFs. Meanwhile, clock showed relatively increase in the expression level with altered 24 hr variation. Bmal1<sup>−/−</sup> MEFs followed the same behavior as Bmal1<sup>−/−</sup> MEFs after transfection with both nCre and cCre fragments. There was no significant difference or effect on the shape or the growth rate between Bmal1<sup>−/−</sup> and Bmal1<sup>fl/fl</sup> MEFs transfected by both nCre and cCre or unsplit-iCRE (Figure 6). CLOCK was translocated into the cytoplasm in both Bmal1<sup>−/−</sup> and Bmal1<sup>fl/fl</sup> MEFs after transfection with both nCre and cCre. This result went along with the previously published data [33] (Supplemental Data S1). This result indicates that this system efficiently succeeded to knock out Bmal1 in MEFs.

Supplemental Data S1: Relative Cre recombinase activity restored on CLOCK localization in Bmal1<sup>−/−</sup> MEF cells.

Western blotting of CLOCK in primary MEF cells of Wild-type, Bmal1<sup>−/−</sup> and Bmal1<sup>−/−</sup> mice. No CLOCK-specific immunoreactivity could be detected in Bmal1<sup>−/−</sup> MEFs. Histone
H4 and β-actin were used as a control for normalizing the nuclear and cytoplasm fractions respectively.

Figure 6: Relative Cre recombinase activity restored on in Bmal1fl/fl MEF cells.

We test the restored recombinase activity using DEX stimulated Bmal1fl/fl MEFs and compared with Bmal1−/− MEFs. DEX stimulated Bmal1−/− MEFs showed decrease in the intensity of the expression level of per1, per2, cry2, dec1 and rev-erba with damped rhythmic expression or totally absent of the 24 hr variation (closed circles) than Wild-type MEFs (open circles). Meanwhile, clock showed relatively increase in the expression level with altered 24 hr variation. Bmal1fl/fl MEFs (closed black circles) followed the same behavior as Bmal1−/− MEFs (open red circles) after transfection with both nCre and cCre. Each value represents the mean +/- S.E. N=3-5.

Discussion

The main conclusion of this study is that our novel split-iCRE complementation system can introduce both temporal and special control of site specific recombination. Split-iCRE system can solve many drawbacks have emerged during the extensive use of unsplit-iCRE recombinase in molecular biology. It has a special importance in neuroscience lacking selective promoter region for conditional knocking out specific gene at specific brain tissue. Many websites, such as Gene Paint; a digital atlas for gene expression pattern in mouse embryo (http://www.genepaint.org/frameset.html) [40], MET scout (http://www.metscout.mpg.de/) [41], GENESAT Brain Atlas of Gene Expression (http://www.gensat.org/cre.jsp) and Brain Stars (http://brainstars.org/) [42] have been emerged. We can determine the suitable promoters that are both expressed in the required tissue with the desired expression intensity. Each cell must express both fragments to have the desired recombination activity. Recently, usage of unsplit-iCRE complementation has been extensively used by several laboratories (both in vitro and in vivo) [22-26]. It has advantages over usage of other solutions for temporal control of the recombinase activity. For example, it is preceded by genetic marking of cells expressing one fragment of split-iCRE (either nCre or cCre), by co-expressing EGFP. Afterwards, we introduce of the other split-iCRE fragment, to obtain the recombinase activity. This procedure gives a chance for direct comparison between the cellular morphology and dynamic behavior before and after genetic modification [43]. Split-iCRE fragment complementation can be also used in ectopic gene expression and cell lineage analyses. It could also allow for detailed analysis of the mutant cell behaviors [44]. Previous trials divided unsplit-iCRE into two fragments such as (nCre; 19-59, cCre; 60-343) [19, 22] and (nCre; 1-190, cCre; 191-343) [26]. They could restore 2-68% of the unsplit-iCRE recombinase activity. These can be due to different reasons. Authors choose the site of splitting relying on the similarity between the amino acid sequence of unsplit-iCRE and Topoisomerase I. Topoisomerase I is cleaved by chymotrypsin into two physically associated fragments [45]. Splitting unsplit-iCRE should consider the difference in the mode of action between these two enzymes [38]. Other depended on the intramolecular interaction among amino acids. Presence of
glycine residues between the two β-sheets could tolerate the addition of extra peptides such as linker and α-helix domain [46]. They used the semi-rigid linker used by Jullien et al. [22, 34]. We designed the linker based on the designing criteria of the linker for efficiently separating domains in bifunctional proteins [47]. Previous trials used rapamycin-dependent dimerization or α-helix interactions for enhancing posttranslational association between the two fragments. Some trials used artificially designed antiparallel leucine zipper to assist protein fragment reconstitution [48]. It is unclear whether it can interact with other cellular proteins containing leucine zipper such as myc. In our study, we used yeast GCN4 coil/coil leucine zipper domain, which previously used by some split-iCRE systems [22]. This leucine zipper is proved to cause no interference with the normal cellular physiology [35]. Split-iCRE (either nCre or cCre) behaves as an independent protein. Each has different kinetics than the other fragment. We tested the recombinase activity on synthetic construct and on mammalian genome. We have optimized the system to choose the best spatial and temporal unsplit-iCRE was not present in all cells. We had to sort cells expressed both nCre and cCre fragments. There are two possibilities for this finding. First, the reconstituted split-iCRE may be too low to induce recombination. Second, there may be difference in the expression pattern between the two promoters due to epigenetic factors.

Tyrosine site-specific recombinases and Topoisomerase IB are characterized by a catalytic 5 amino acids residue, Arg-Lys-(His/Lys)-Arg-(His/Trp). This residue assists cleavage and joining steps of DNA recombination and relaxation, respectively [49]. In Cre Recombinase, these residues are Arg-173, Lys-201, His-289, Arg-292 and Trp-315 [49, 50]. So, we have to keep these residues intact and in the same direction to avoid losing the enzymatic activity. Tyr-324 is a nucleophilic attack for the scissile phosphodiester bond in the DNA strand forming a 3′ covalent adduct with the cleaved strand. The formed 5′-hydroxyl group at the cleavage step carries out another nucleophilic attack on the tyrosyl intermediate to promote strand joining. A recombination event is completed in two steps of concerted single strand exchanges. The first gives rise to a Holliday junction intermediate; the second resolves it into reciprocal recombinant products [29, 38, 39]. We choose point of splitting Lys130 and Asp132. This position is within the two interhelical loops of the unsplit-iCRE. It is fully exposed upward to the external environment. Moreover, it is limitedly involved in intra- or intermolecular interactions. Splitting at this site preserves 3D-configuration of subdomains that remain structured in each isolated fragment. Meanwhile, our control was the unsplit-iCRE. We have missing Glutamine 131 from unsplit-iCRE. The resultant complemented split-iCRE lacked the recombinase activity (Data is not shown). This may be due to conformational changes by eliminating this amino acid hindering the interaction between the C-clamp shaped loop and the loxP site. There are some α-helix domains within the unsplit-iCRE. So, we expect that there are transient interactions among them for each preserved split-iCRE fragment. To regulate the interaction between fragments, we relied on the heterodimerization of yeast derived leucine zipper moieties. Interaction between these leucine zipper domains induced complementation and restoring the unsplit-iCRE recombinase activity. To avoid the major steric hindrance, we used long rigid linker to separate between the two active moieties and we deleted one amino acid; Glutamine 131 to give space (~13.89Å) for best fitting between the reconstituted Cre and DNA. We tested the dynamics of each protein fragment separately. cCre fragment might be less stable than nCre fragment. So, it is recommended to control the cCre expression by stronger promoter. Addition of external NLS, enhances the restored Cre recombinase activity. Although we did not directly compare among our constructs and other constructs, we could compare the published data for another variant of the steroid-regulated systems (RU-486 and tamoxifen) [18-21], or other fragments [22, 25, 26]. This comparison suggests that our novel system has a higher regulated induction ratio than these systems. Interestingly, split-iCRE system proved recently its efficiency for modulation of recombinase activity in prokaryotic cells as well [44, 45].
The split-iCRE system shows higher sensitivity to the dimerizers than other splitted functional proteins. Moreover, there is a significant shift of the dose-response curves to the left relative to published data [50, 51]. However, at a relatively low concentration of active unsplit-iCRE inside the nucleus recombination can occur [52]. Thus, we added NLS to both nCre and cCre fragments to increase the concentration of complemented split-iCRE moieties in the nucleus. So, even at a level below the Kd value of association of yeast GCN4 coil/coil, it will result in an active complemented split-iCRE sufficient for recombination, even if the proportion of this complex relative to the total concentration of unsplit-iCRE moieties, as described by the classical binding kinetics, is low. But, such case may induce disturbance of the mammalian genome due to the non-specific recombination with loxP-like sequence. It is likely that cCre can bind to loxP sequence, similar to the unsplit-iCRE (119-343) [38]. However, the correct positioning and/or stabilization of the right conformation of cCre by the nCre seem to be required to achieve the recombination. In the unsplit-iCRE, helices B and D contact the loxP site, while the loop A/B is engaged in interactions with helix E of the same or adjacent Cre [5-8, 38].

It can be supposed that these interactions also occur for the isolated nCre and cCre. Moreover, given the interactions of helices A and B with helices C and D, it is likely that the isolated nCre possesses a structure close to the one it adopts within the unsplit-iCRE, and that would also favor its ability to stabilize the complex. The appearance of recombinase activity would then be the consequence of the stepwise assembly, on the loxP site, of a native-like conformation of unsplit-iCRE by the isolated moieties, involving, first, the binding of nCre to the loxP half site, followed by the binding of cCre. The subsequent stabilization of the ensemble through interactions of nCre with cCre, and also with nCre moieties fixed to the adjacent loxP half site. On the other hand, their strong association induced by the dimerizer compensates for the first effect. It also potentiates the interactions between the nCre and cCre moieties. This will help in the emergence of the correct conformation of the assembly and, through that, of enzymatic activity. Previously used split-iCRE system (nCre; 19-59 and cCre; 60-343) may have the same mechanism. However, the interactions of the nCre with both the rest of the molecule and with loxP are rare and weaker than those of nCre of our system. Moreover, the structure of previously used nCre reproduced less well the structure it adopts within the intact unsplit-iCRE. In fact, we can assume that the two α-helices do not form their V-like native structure. Moreover, in the previous system, concurrent binding of nCre to loxP and cCre bound to the same site will lead it to adopt unsplit-iCRE-like conformation. That, in turn, will stabilize the whole complex. This mechanism of action, i.e. interaction-induced folding of proteins from complementary peptide fragments, has been invoked to explain the protein fragment implementation assay [53] that presents similarities with our approach. Yet, this mechanism was postulated to work only for monomeric enzymes, while the present example suggests that it might also work for more complex, multimeric, enzymes (in Cre-loxP interaction, we need four molecules of unsplit-iCRE for the DNA recombination). However, in a stable expression context, the limiting factor seems not to be the intrinsic activity, presumably because of the accumulation of enough protein in the cell to compensate for the differences in the intrinsic activities. The same conclusion can be drawn from the global comparison of CrePR1 and the different unsplit-iCRE variants. The prime criterion should rather be the background activity of the system, and from this point of view the unsplit-iCRE system based on nCre and cCre clearly performs better, as not only does it achieve full recombination but it also possesses a low background activity. Thus, it should represent a good solution when a tight temporal control of recombinase activity is required, as in the case of the creation of conditional knock-out animals. Further investigations are required for examining the applicability of this system in vivo. Finally, our results show that the controlled dimerization of totally inactive fragments may represent an efficient technique for the regulation, within the cell, of the activity of any functional protein. It can be used even if the enzyme has a relatively complex mechanism of action. Thus, it may complement already existing approaches, such as the regulation of expression through modified promoters.
or the use of compounds interacting specifically with this enzyme.

Acknowledgments
We are indebted to Dr. Gregor Eichele (Professor, Department of Genes and Behavior, Max-Planck Institute for Biophysical Chemistry, Göttingen, Germany) for funding part of the research on this project. We are grateful to Dr. Henrik Oster (Professor of Chronophysicsiology at the Medical Department I, University of Lübeck, Germany) for providing part of the research on this project. We are grateful to Dr. Johannes Hirrlinger (Professor, Faculty of Medicine, Interdisciplinary Centre for Clinical Research, University of Leipzig, Germany) for providing yeast GCN4 coil/coil plasmid vector. We thank Dr. Xunlei Zhou, University of Heidelberg, Germany) for providing us with the humanized version of unsplit-iCRE recombinase using the pNLS-iCre cDNA (with nuclear localization signal present in the Cre’s C-terminus. We thank Dr. Satoru Koyanagi (Professor, Department of Pharmaceutics, Faculty of Pharmaceutical Sciences, Kyushu University, Fukuoka, Japan) for revising this manuscript.

References


